UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/779,402	02/13/2004	Steven J. McCarthy	ID-504 (80226)	2799

89137 7590 10/06/2011

ADDMG - RIM
255 S. Orange Avenue
Suite 1401
Orlando, FL 32801

EXAMINER
NGUYEN, MINH CHAU

ART UNIT	PAPER NUMBER
2442	

NOTIFICATION DATE	DELIVERY MODE
10/06/2011	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

creganoa@addmg.com
portfolioprosecution@rim.com

UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

Application Number: 10/779,402
Filing Date: February 13, 2004
Appellant(s): MCCARTHY ET AL.

Jeremy B. Berman, ESQ.
(Reg. No.: 60,582)
For Appellant

EXAMINER'S ANSWER

This is in response to the appeal brief filed July 27, 2011 appealing from the Office action mailed February 17, 2011.

(1) Real Party in Interest

The examiner has no comment on the statement, or lack of statement, identifying by name the real party in interest in the brief.

(2) Related Appeals and Interferences

The following are the related appeals, interferences, and judicial proceedings known to the examiner which may be related to, directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal.

(3) Status of Claims

The following is a list of claims that are rejected and pending in the application:

Claims 1-3, 5-6, 8-10, 12, 14-15, 17-18, 20 and 22-28.

(4) Status of Amendments After Final

The examiner has no comment on the appellant's statement of the status of amendments after final rejection contained in the brief.

(5) Summary of Claimed Subject Matter

The examiner has no comment on the summary of claimed subject matter contained in the brief.

(6) Grounds of Rejection to be Reviewed on Appeal

The examiner has no comment on the appellant's statement of the grounds of rejection to be reviewed on appeal. Every ground of rejection set forth in the Office action from which the appeal is taken (as modified by any advisory actions) is being maintained by the examiner except for the grounds of rejection (if any) listed under the

subheading "WITHDRAWN REJECTIONS." New grounds of rejection (if any) are provided under the subheading "NEW GROUNDS OF REJECTION."

(7) Claims Appendix

The examiner has no comment on the copy of the appealed claims contained in the Appendix to the appellant's brief.

(8) Evidence Relied Upon

6,970,913	Albert et al.	11-2005
2004/0205120	Dar et al.	10-2004

(9) Grounds of Rejection

The following ground(s) of rejection are applicable to the appealed claims:

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 1-3, 5-6, 8-10, 12, 14-15, 17-18, 20 and 22-28 are rejected under 35 U.S.C. 103(a) as being unpatentable over Albert et al. (US 6,970,913) hereinafter "**Albert**", and in view of Dar et al. (US 2004/0205120) hereinafter "**Dar**".

Claim 1

Albert teaches a communications system comprising:

a plurality of servers [*i.e. servers 221-223*] connected together in a network [*i.e. network 210*] for processing a plurality of different job types [*i.e. processing a plurality of task types for flowing of packets or handling connections (e.g. "routing the packet, gathering statistics...modifying packet" in col. 7, lines 14-18; and "Forwarding agents can accomplish their required tasks" in col. 8, line 11); and these task types will be implemented by a plurality of virtual machines on the servers*] having respective resource usage [*i.e. processing capacity usage*] associated therewith (Albert, figure 2A; col. 6, line 51-col. 7, line 30; col. 8, lines 8-12, lines 57-67; col. 28, lines 9-47; col. 29, line 44-col. 30, line 49; col. 32, lines 14-51);

each server, after beginning execution of at least one job, determining a respective health metric thereof based upon at least one job being executed thereby [*i.e. "The nature of the feedback messages from the real machines is that the messages somehow express the level of load on the real machine as a result of handling connections" and "the server determines the usage of processing capacity for each of the virtual machines that is being implemented" in col. 30, lines 1-3, lines 21-23, thus the usage of processing capacity is considered as the health metric which determined from the result after the task being executed*] and weighting the health metric of the at least one job [*i.e. weighting the usage of processing capacity to obtain "a weight" for the virtual machine*] (Albert, col. 28, lines 9-47; col. 29, line 44-col. 30, line 49; col. 32, lines 14-51); and

said servers mapping the weighted health metrics to a common scale [*i.e. expressing the weights to a common level (e.g. high or low level)*] (Albert, col. 30, line 1-col. 31, line 3; col. 32, lines 14-51); and

a dispatcher [*i.e. service manager 1 or 1140*] for collecting the commonly scaled weighted health metrics [*i.e. weights or levels of load*] from said servers [*i.e. the servers the real machines*] by polling said servers for the weighted health metrics [*i.e. retrieving is considered as polling the weights of the real machines*] and distributing jobs to said servers based thereon [*i.e. assigning connection tasks to the real machines based on the weights*] (Albert, figures 2A, 11A; col. 6, line 51-col. 7, line 30; col. 8, lines 8-12, lines 57-67; col. 28, lines 9-47; col. 29, line 44-col. 30, line 49; col. 32, lines 14-51).

Albert fails to teach different resource usage characteristics; and the health metric is determined based upon the respective resource usage characteristic after execution of the job, the resource usage characteristic representing resources being consumed by the at least one job; and the weighted health metrics for different resource usage characteristics.

However, in an analogous art, **Dar** teaches different resource usage characteristics [*i.e. processor (CPU), memory, and input/ouput (I/O) metrics*] (Dar, paragraphs 0029); the health metric is determined based upon the respective resource usage characteristic after execution of the job [*i.e. the health of the server is considered as the heath metric which aggregating the metrics.*

Moreover, Dar discloses "This monitoring may be periodic, e.g., every 10 seconds...or a synchronous monitoring of a different period would be acceptable", thus the monitoring could be run for obtaining the health and/or the metrics after a period of time of the program execution], the resource usage characteristic representing resources being consumed by the at least one job [*i.e. the metrics are consumed by the program*] (Dar, paragraphs 0001, 0027, 0029-0032); and the weighted health metrics for different resource usage characteristics [*i.e. the total loads for the servers for different resource (e.g. CPU, memory, I/O) loads*] (Dar, paragraphs 0001, 0027, 0029-0032).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to incorporate the features of different resource usage characteristics; and the health metric is determined based upon the respective resource usage characteristic after execution of the job, the resource usage characteristic representing resources being consumed by the at least one job; and the weighted health metrics for different resource usage characteristics, as disclosed by Dar, into the teachings of Albert. One would be motivated to provide “one or more of the following advantages...to improve resource utilization by the servers...Network server load balancing may be improved. Availability and/or scalability of network servers can be improved" [Dar, paragraph 0016].

Claim 2

Albert in combination with Dar teach the communications system of Claim 1 wherein the resource usage characteristics comprise at least one processing utilization characteristic and at least one input/output utilization characteristic [*i.e. processor (CPU) and input/ouput (I/O) metrics*] (Dar, paragraph 0029).

Claim 3

Albert in combination with Dar teach the communications system of Claim 1 further comprising a knowledge base [*i.e. a database*] for cooperating with said dispatcher [*i.e. the service manager*] for storing the weighted health metrics [*i.e. storing the weights*] (Albert, col. 31, lines 49-59).

Claim 5

Albert in combination with Dar teach the communications system of Claim 1 wherein said servers provide completed job results to said dispatcher, and wherein the weighted health metrics are provided to said dispatcher with the completed job results [*i.e. "the feedback messages from the real machines is that the messages somehow express the level of load on the real machine as a result of handling connections", and "a process executed on a server for determining a weight to be sent to the service manager in a feedback message… to determining the weights..., in a step 1206, the*

server determines the remaining processing capacity", thus the weights are sent to the service manger for the completed job/process results] (Albert, Col. 30, L. 1-49).

Claim 6

Albert in combination with Dar teach the communications system of Claim 1 further comprising at least one load generator [*i.e. load balancer*] for generating the jobs for said servers and communicating the jobs to said dispatcher; and wherein said dispatcher [*i.e. the service manager*] further provides the completed job results to said at least one load generator [*i.e. "load balancer to be located between group of servers. It would be advantageous if a distributed architecture could be used for load balancing and if a server manager could be used to gather feedback from servers to make better load balancing decisions" in col. 3, lines 60-64*] (Albert, col. 3, line 59-col. 4, line 3; col. 8, lines 57-67; col. 11, lines 56-65; col. 28, lines 9-47; col. 30, lines 1-49; col. 32, lines 14-51).

Claim 8

Albert in combination with Dar teach the communications system of Claim 1 wherein the jobs relate to electronic mail (e-mail) processing [*i.e. the programs provide email service*] (Dar, paragraph 0001). One would be motivated to provide a wide array of services to clients via the network [Dar, paragraph 0001]

Claim 9

Albert teaches a load distributor for a plurality of servers [*i.e. servers 221-223*] connected together in a network [*i.e. network 210*] for processing a plurality of different job types [*i.e. processing a plurality of task types for flowing of packets or handling connections (e.g. "routing the packet, gathering statistics...modifying packet" in col. 7, lines 14-18; and "Forwarding agents can accomplish their required tasks" in col. 8, line 11); and these task types will be implemented by a plurality of virtual machines on the servers*] having respective resource usage [*i.e. processing capacity usage*] associated therewith (Albert, figure 2A; col. 6, line 51-col. 7, line 30; col. 8, lines 8-12, lines 57-67; col. 28, lines 9-47; col. 29, line 44-col. 30, line 49; col. 32, lines 14-51); each server, after beginning execution of at least one job, determining a respective health metric thereof based upon at least one job being executed thereby [*i.e. "The nature of the feedback messages from the real machines is that the messages somehow express the level of load on the real machine as a result of handling connections" and "the server determines the usage of processing capacity for each of the virtual machines that is being implemented" in col. 30, lines 1-3, lines 21-23, thus the usage of processing capacity is considered as the health metric which determined from the result after the task being executed*] and weighting the health metric of the at least one job [*i.e. weighting the usage of processing capacity to obtain "a weight" for the virtual machine*] (Albert, col. 28, lines 9-47; col. 29, line 44-col. 30, line 49; col. 32, lines 14-51), the load distributor comprising:

a dispatcher [*i.e. service manager 1 or 1140*] for collecting the commonly scaled weighted health metrics [*i.e. weights or levels of load*] from said servers [*i.e. the servers or the real machines*] by polling said servers for the weighted health metrics [*i.e. retrieving is considered as polling the weights of the real machines*] and distributing jobs to said servers based thereon [*i.e. assigning connection tasks to the real machines based on the weights*] (Albert, figures 2A, 11A; col. 6, line 51-col. 7, line 30; col. 8, lines 8-12, lines 57-67; col. 28, lines 9-47; col. 29, line 44-col. 30, line 49; col. 32, lines 14-51); and

said servers mapping the weighted health metrics to a common scale [*i.e. expressing the weights to a common level (e.g. high or low level)*] (Albert, col. 30, line 1-col. 31, line 3; col. 32, lines 14-51); and

a knowledge base [*i.e. a database*] for cooperating with said dispatcher [*i.e. the service manager*] for storing the weighted health metrics [*i.e. storing the weights*] (Albert, col. 31, lines 49-59).

Albert fails to teach different resource usage characteristics; and the health metric is determined based upon the respective resource usage characteristic after execution of the job, the resource usage characteristic representing resources being consumed by the at least one job; and the weighted health metrics for different resource usage characteristics.

However, in an analogous art, **Dar** teaches different resource usage characteristics [*i.e. processor (CPU), memory, and input/ouput (I/O) metrics*]

(Dar, paragraphs 0029); the health metric is determined based upon the respective resource usage characteristic after execution of the job [*i.e. the health of the server is considered as the heath metric which aggregating the metrics. Moreover, Dar discloses "This monitoring may be periodic, e.g., every 10 seconds...or a synchronous monitoring of a different period would be acceptable", thus the monitoring could be run for obtaining the health and/or the metrics after a period of time of the program execution*], the resource usage characteristic representing resources being consumed by the at least one job [*i.e. the metrics are consumed by the program*] (Dar, paragraphs 0001, 0027, 0029-0032); and the weighted health metrics for different resource usage characteristics [*i.e. the total loads for the servers for different resource (e.g. CPU, memory, I/O) loads*] (Dar, paragraphs 0001, 0027, 0029-0032).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to incorporate the features of different resource usage characteristics; and the health metric is determined based upon the respective resource usage characteristic after execution of the job, the resource usage characteristic representing resources being consumed by the at least one job; and the weighted health metrics for different resource usage characteristics, as disclosed by Dar, into the teachings of Albert. One would be motivated to provide “one or more of the following advantages...to improve resource utilization by the servers...Network server load balancing may be

improved. Availability and/or scalability of network servers can be improved" [Dar, paragraph 0016].

Claim 14

Albert teaches a job distribution method for a plurality of servers a plurality of servers [*i.e. servers 221-223*] connected together in a network [*i.e. network 210*], the servers for processing a plurality of different job types [*i.e. processing a plurality of task types for flowing of packets or handling connections (e.g. "routing the packet, gathering statistics...modifying packet" in col. 7, lines 14-18; and "Forwarding agents can accomplish their required tasks" in col. 8, line 11); and these task types will be implemented by a plurality of virtual machines on the servers*] having respective resource usage [*i.e. processing capacity usage*] associated therewith (Albert, figure 2A; col. 6, line 51-col. 7, line 30; col. 8, lines 8-12, lines 57-67; col. 28, lines 9-47; col. 29, line 44-col. 30, line 49; col. 32, lines 14-51), the method comprising:

determining a respective health metric of each server after it begins execution of at least one job based upon the at least one job being executed thereby [*i.e. "The nature of the feedback messages from the real machines is that the messages somehow express the level of load on the real machine as a result of handling connections" and "the server determines the usage of processing capacity for each of the virtual machines that is being implemented" in col. 30, lines 1-3, lines 21-23, thus the level of load of each real machine is*

determined based on the usage of processing capacity from the result after the task being executed] and weighting the health metric of the at least one job [*i.e. weighting the usage of processing capacity to obtain "a weight"*] (Albert, col. 28, lines 9-47; col. 29, line 44-col. 30, line 49; col. 32, lines 14-51);

polling said servers for the weighted health metrics [*i.e. retrieving is considered as polling the weights of the real machines*] (Albert, col. 32, lines 14-51) and mapping the weighted health metrics to a common scale [*i.e. expressing the weights to a common level (e.g. high or low level)*] (Albert, col. 30, line 1-col. 31, line 3; col. 32, lines 14-51); and

distributing jobs to said servers based upon the commonly scaled weighted health metrics [*i.e. assigning connection tasks to the real machines based on the weights*] (Albert, col. 28, lines 9-47; col. 29, line 44-col. 30, line 49; col. 32, lines 14-51).

Albert fails to teach different resource usage characteristics; and the health metric is determined based upon the respective resource usage characteristic after execution of the job, the resource usage characteristic representing resources being consumed by the at least one job; and the weighted health metrics for different resource usage characteristics.

However, in an analogous art, **Dar** teaches different resource usage characteristics [*i.e. processor (CPU), memory, and input/ouput (I/O) metrics*] (Dar, paragraphs 0029); the health metric is determined based upon the

respective resource usage characteristic after execution of the job [*i.e. the health of the server is considered as the heath metric which aggregating the metrics. Moreover, Dar discloses "This monitoring may be periodic, e.g., every 10 seconds...or a synchronous monitoring of a different period would be acceptable", thus the monitoring could be run for obtaining the health and/or the metrics after a period of time of the program execution*], the resource usage characteristic representing resources being consumed by the at least one job [*i.e. the metrics are consumed by the program*] (Dar, paragraphs 0001, 0027, 0029-0032); and the weighted health metrics for different resource usage characteristics [*i.e. the total loads for the servers for different resource (e.g. CPU, memory, I/O) loads*] (Dar, paragraphs 0001, 0027, 0029-0032).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to incorporate the features of different resource usage characteristics; and the health metric is determined based upon the respective resource usage characteristic after execution of the job, the resource usage characteristic representing resources being consumed by the at least one job; and the weighted health metrics for different resource usage characteristics, as disclosed by Dar, into the teachings of Albert. One would be motivated to provide “one or more of the following advantages...to improve resource utilization by the servers...Network server load balancing may be improved. Availability and/or scalability of network servers can be improved" [Dar, paragraph 0016].

Claim 22

Albert in combination with Dar teach the communications system of Claim 1, wherein the at least one job comprises e-mail delivery [*i.e. the programs provide email service, and the email delivery is included in the email service*] (Dar, paragraph 0001). One would be motivated to provide a wide array of services to clients via the network [Dar, paragraph 0001].

Claim 23

Albert in combination with Dar teach the communications system of Claim 6, wherein said at least one load generator [*i.e. load balancing]* comprises an e-mail aggregation engine [*i.e. the programs provide email service, and the load balancing between the servers for performing the programs based on the aggregating metrics, thus the examiner interprets that the aggregating metrics would be done by an aggregation engine*] (Dar, paragraphs 0001, 0016, 0029, 0033).

Claim 24

Albert in combination with Dar teach the communications system of Claim 6, wherein said servers also provide completed job results to said at least one load

generator [*i.e. "load balancer to be located between group of servers. It would be advantageous if a distributed architecture could be used for load balancing and if a server manager could be used to gather feedback from servers to make better load balancing decisions" in col. 3, lines 60-64*] (Albert, col. 3, line 59-col. 4, line 3; col. 8, lines 57-67; col. 11, lines 56-65; col. 28, lines 9-47; col. 30, lines 1-49; col. 32, lines 14-51).

Claims 10, 12, 25, 26 are corresponding apparatus claims of system claims 2, 5, 22, 6. Therefore, they are rejected under the same rationale.

Claims 15, 27 are corresponding method claims of system claims 2, 22. Therefore, they are rejected under the same rationale.

Claims 17-18, 20, 28 are corresponding computer-readable medium claims of apparatus claims 9-10, 12, 25. Therefore, they are rejected under the same rationale.

(10) Response to Argument

Argument 1:

Regarding to claim 1, "*Application respectfully submits that the Examiner's interpretation of the prior art is flawed, however. The very portion of Albert et al. cited by the Examiner explains that the feedback messages may be sent to the*

service manager from either an individual server, or from a representative server that collects server feedback messages. This is not polling. As understood by those of skill in the art, a service manager polling a server for data would mean that the service manager repeatedly asks the server for the data, and the server sends the data in response when it has the data to send. In sharp contrast to this commonly understood definition of polling, the servers of Albert et al. send their feedback messages to the service manager without any request therefrom. The service manager of Albert et al. is therefore not polling the servers for their feedback messages.

Consequently, Albert et al., and therefore the combination of Albert et al. and Dar, fails to disclose a dispatcher for collecting the commonly scaled weighted health metrics from the servers by polling the servers for the weighted health metrics and distributing jobs to the servers based thereon, as recited in independent claim 1" [See Appeal Brief, (7) Argument, pages 8-9].

In response to Appellant's argument, the Examiner asserts that:

Albert et al. discloses "*each real machine also includes an interface that allows it send feedback information to service manager that tells service manager 1140 how busy the real machine is handling connections for various virtual machines…Each real machine may have its own connection to service manager 1140 and may independently send feedback messages to service manger 1140*" (see Albert, col. 29, lines 25-39); and "*each real machine independently sends messages to service manager 1140 over a feeback message bus. Real machines 1102, 1104, and 1106 each are connected to a*

feedback message bus 1110 that is connected to service manager 1140" (see Albert, col. 29, lines 44-48 and figure 11B); and "*Service manager 1140 includes an interface that receives the feedback messages from the real machines*" (see col. 29, lines 63-65) ; and "*When the service manager receives a feedback packet from a server, the service manager updates a database of available hosts for each different virtual address, port and protocol that is being supported by the service manager. FIG. 14 is a diagram illustrating a data structure 1400 for a first virutal machine port and protocol that includes all of the servers used to implement that virtual machine port and protocol along with a weight for each server. The weight is received as part of a feedback packet. A second data structure 1402 stores the identity of the servers and their respective weights for a second virtual machine port and protocol…a process implemented on a service manager for updating the data structures illustrated in FIG. 14 upon receipt of a feedback packet…, the service manager receives an update message from a server…*" (see Albert, col. 31, lines 49-67); and "*the service manager receives feedback packets from each server that include weights for the server that control the selection of that server for handling connections to a virtual machine implemented on the server…the service manager retrieves the real machine weights for the virtual machine. The real machine weights represent the capacity of each real machine to handle connections for that specific virtual machine…The real machine is selected by the service manager using the weights retrieved from its database for the virtual machine*" (see Albert, col. 32, lines 14-40).

According to these citations, the Exmainer interprets that the service manager receives the weights of the real machines (or servers), which are included in the feedback messages, from each real machine or server independently. While the service manager receives the feedback messages from the real machines, the weights data are stored in the database, and the database is a data structure to kept or collect the weights data. Then, the service manager retrieves the real machines' weights from the database, so that the retrieved real machines' weights are actually received from the real machines (or servers) via the feedback messages. Thus, the Examiner interprets that retrieving the real machines' weights, which are received from the real machines (or servers), is considered as polling the servers' weighted health metrics.

Therefore, the Examiner assets that Albert's sevice manager for storing and retrieving the weights that sent from the real machines via the feedback messages, and also for selecting real machine(s) to handle connections based on the retrieved weights to read on Appellant's dispatcher for collecting the commonly scaled weighted health metrics from the servers by polling the servers for the weighted health metrics and distributing jobs to the servers based thereon.

In conclusion, the combination of Albert et al. and Dar, has not failed to disclose *"a dispatcher for collecting the commonly scaled weighted health metrics from the servers by polling the servers for the weighted health metrics and distributing jobs to the servers based thereon"* as recited in claim 1.

Regarding to claims 9, 14 and 17 that contain similar recitations to claim 1.

Therefore, the Exmainer submits the same response as that for claims 9, 14 and 17.

(11) Related Proceeding(s) Appendix

No decision rendered by a court or the Board is identified by the examiner in the Related Appeals and Interferences section of this examiner's answer.

For the above reasons, it is believed that the rejections should be sustained.

Respectfully submitted,

/M. N./

Examiner: Minh-Chau Nguyen, Art Unit 2442

/John Follansbee/

Supervisory Patent Examiner, Art Unit 2451

September 26, 2011

Conferees:

/KEVIN BATES/
Primary Examiner, Art Unit 2456

/John Follansbee/
Supervisory Patent Examiner, Art Unit 2451